ELSEVIER

Journal of Economic Behavior & Organization
Vol. 65 (2008) 436–457

JOURNAL OF
Economic Behavior
& Organization

www.elsevier.com/locate/econbase

Heterogeneous social preferences

Mathias Erlei

Institute of Business Administration and Economics, Clausthal University of Technology, Julius-Albert-Str. 2, 38678 Clausthal-Zellerfeld, Germany

Received 28 June 2004; accepted 16 February 2006
Available online 13 November 2006

Abstract

Recent research has shown the usefulness of social preferences in explaining behavior in laboratory experiments. This paper demonstrates that models of social preferences are particularly powerful in explaining behavior if they are embedded in a setting of heterogeneous actors with heterogeneous (social) preferences. For this purpose a simple model is introduced that combines the basic ideas of inequity aversion, social welfare preferences, reciprocity and heterogeneity. This model is applied to 43 games, and its predictive accuracy is clearly higher than that of the isolated approaches. Furthermore, it can explain most of the "anomalies" discussed in Goeree and Holt [Goeree, J., Holt, Ch.A., 2001. Ten little treasures of game theory and ten intuitive contradictions, American Economic Review 91, 1402–1422].

JEL classification: C72; C92; D63

Keywords: Social preferences; Reciprocity; Heterogeneity; Inequity aversion; Welfare Preferences

1. Introduction

When we meet somebody for the first time we often ask ourselves what *kind* of person he or she is. This is not just a matter of pure curiosity; it is mostly *the* central question in deciding whether to have further contact with that person. Such a way of thinking suggests that people are indeed very different from each other and that the type of the person we are interacting with is of utmost importance. However, most models in economic theory and particularly in game theory assume homogeneous actors with identical preferences. Differences in behavior usually stem from different initial equipments and different positions in the game under consideration. The main purpose of this paper is to demonstrate the usefulness of explicitly modeling heterogeneity

E-mail address: m.erlei@tu-clausthal.de.

0167-2681/$ – see front matter
doi:10.1016/j.jebo.2006.02.006

of preferences in explaining the behavior of subjects in 43 laboratory experiments. We also try to corroborate the relevance of social preferences in explaining human behavior.

Experimental evidence shows clearly that there are many games in which Standard Nash Equilibrium[1] (SNE) describes people's behavior quite well. However, there seem to be just as many other games in which laboratory behavior deviates substantially from the predictions of standard game theory. Obviously, there is a need for theoretical innovations that can explain the successes of game theory as well as its failures. Without doubt, theory has reacted to experimental evidence. There are several new theoretical approaches that can claim to have at least partial success in introducing superior concepts. Dynamic evolutionary approaches[2] (e.g. replicator dynamics) often but not always converge to SNE. Quantal Response Equilibria and in particular the Logit Equilibrium (McKelvey and Palfrey, 1995, 1998; Goeree and Holt, 2001) have been quite successful in explaining behavioral reactions due to parameter variations in games with identical SNEs. Finally, there is a third strand of research that was successful in explaining deviations from SNE. These are approaches of social preferences. The social preferences approach can be divided into at least three important substrands: theories of intentional reciprocity (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004), the inequity aversion approach (Bolton and Ockenfels, 2000; Fehr and Schmidt, 1999) and recently a theory of social welfare preferences (Andreoni and Miller, 2002; Charness and Rabin, 2002). In this paper we shall concentrate on the last two approaches.

Bolton and Ockenfels as well as Fehr and Schmidt introduce concepts of inequity aversion. It is assumed that there exist people who dislike inequality and who actually sacrifice money to reduce it. Both concepts are particularly successful in describing laboratory behavior when they assume heterogeneous actors. Bolton and Ockenfels' model is exclusively defined for heterogeneous subjects. Although Fehr and Schmidt's model can be used for homogeneous populations, all successful applications assume a mixture of inequity averse and strictly egoistic subjects; the latter being individuals with the standard utility functions in game theory. The approaches differ in the concrete definition of inequity aversion, and Bolton and Ockenfels allow for more general preference distributions of subjects. However, the general version of their model is somewhat more complicated, making it less suitable for direct application. It is no surprise that most further applications of the inequity aversion approach use the simpler Fehr and Schmidt variant. In the meantime inequity aversion has been challenged by numerous experiments (e.g. Kagel and Wolfe, 2001; Charness and Rabin, 2002; Engelmann and Strobel, 2004).

The most important alternative to inequity aversion has been presented by Charness and Rabin. They introduce a model of social welfare preferences with and without reciprocity. Social welfare preferences are characterized by individuals who give positive weight to aggregated surplus (i.e. if other people are better off, c.p., utility of individuals increase). The authors carried out 32 experiments and compared theoretical predictions of several social preference approaches with the experimental data. They concluded that social welfare preferences provide the best fit to the data. However, the comparison between social welfare preferences and the inequity aversion model is biased because Charness and Rabin ignore the fact that the most fruitful version of Fehr and Schmidt's inequity aversion model takes explicitly into account that there is a heterogeneity of preferences. In fact, in Fehr et al. (2005) as well as in Fehr et al. (2004) inequity averse actors are only a minority of the population, and the explanatory power of the model stems in particular

[1] Here and in the remaining part of the paper "Standard Nash Equilibrium" means Nash Equilibrium in games among players with the standard assumptions about purely self-centered utility (i.e. players without any kind of social preferences).

[2] See Weibull (1995) or Fudenberg and Levine (1998).